Modelling daily streamflow at ungauged catchments: what information is necessary?

Patil, S.D.; Stieglitz, M.

Hydrological Processes

DOI:
10.1002/hyp.9660

Published: 28/12/2012

Peer reviewed version

Modeling daily streamflow at ungauged catchments: What information is necessary?

Sopan Patil[1*] and Marc Stieglitz[1, 2]

[1] School of Civil and Environmental Engineering,
Georgia Institute of Technology, Atlanta, GA 30332

[2] School of Earth and Atmospheric Sciences,
Georgia Institute of Technology, Atlanta, GA 30332

* Current affiliation: National Health and Environmental Effects Research Laboratory,
U.S. Environmental Protection Agency, Corvallis, OR 97333

Corresponding author: sopan.patil@gmail.com

Submission to:

Hydrological Processes

Abstract

Rainfall-runoff modeling at ungauged catchments often involves the transfer of calibrated model parameters from "donor" gauged catchments. However, in any rainfall-runoff model, some parameters tend to be more sensitive to the objective function, whereas others are insensitive over their entire feasible range. In this paper, we analyze the effect of selectively transferring sensitive vs. insensitive parameters on streamflow predictability at ungauged catchments. We develop a simple daily time-step rainfall-runoff model (EXP-HYDRO) and calibrate it at 756 catchments within the continental United States. Nash Sutcliffe efficiency of $\sqrt{Q}$ (NS) is used as the objective function. The model simulates satisfactorily at 323 catchments (NS > 0.6), most of which are located in the eastern part of US, along the Rocky Mountain Range, and near the western Pacific coast. Of the six calibration parameters, only three parameters are found to be sensitive to NS. Two of these parameters control the hydrograph recession behavior of a catchment and the third parameter controls the snowmelt rate. We find that when only sensitive parameters are transferred, model performance at ungauged catchments is almost on par with that of transferring all six parameters. Conversely, transfer of only insensitive parameters results in significant deterioration of model performance. Results suggest that streamflow predictability at ungauged catchments using rainfall-runoff models is largely dependent on the transfer of a small subset of parameters. We recommend that, in any modeling framework, such parameters should be identified and further characterized to better understand the information controlling streamflow predictability at ungauged catchments.

1. Introduction

Rainfall-runoff models are the essential tools for prediction of catchment streamflow and are applied for numerous tasks in hydrology. These tasks include: short-term streamflow forecasting [*Zealand et al.*, 1999; *Shukla and Lettenmaier*, 2011], flood frequency estimation [*Merz and Blöschl*, 2005; *Moretti and Montanari*, 2008], water quality assessment [*Krysanova et al.*, 1998; *Servais et al.*, 2007], low flow predictions [*Smakhtin*, 2001; *Rees et al.*, 2004; *Staudinger et al.*, 2011], study of the ecosystem services linked to catchment hydrologic functioning [*Poff et al.*, 2010; *Abdelnour et al.*, 2011; *Notter et al.*, 2012], and assessment of climate change impacts on water availability [*Hamlet and Lettenmaier*, 1999; *Christensen et al.*, 2004; *Xu et al.*, 2011]. A variety of rainfall-runoff models have been developed over the years and successfully implemented at catchments across the world (see reviews by *Singh* [1995], *Beven* [2001], *Singh and Woolhiser* [2002], and *Singh and Frevert* [2006]). However, regardless of the model used, an important prerequisite for streamflow prediction involves calibration of model parameters using observed streamflow data [*Beven*, 2001]. Unfortunately, majority of the catchments throughout the world are ungauged (i.e., they lack streamflow observations). Therefore, a challenge for hydrologists is to develop tools and strategies for predicting streamflow at these ungauged catchments [*Sivapalan et al.*, 2003; *Wagener and Montanari*, 2011].

A common strategy for streamflow modeling at ungauged catchments involves the following procedure: (1) calibration of model parameters at gauged catchments using the observed streamflow data, and (2) transfer of the calibrated parameters from gauged to ungauged catchments that are perceived to be hydrologically similar [*Oudin et al.*, 2010]. Here, we define two or more catchments as hydrologically similar if their daily stream responses (runoff) are highly correlated to each other [*Archfield and Vogel*, 2010; *Patil and Stieglitz*, 2012]. Since streamflow data is not available at ungauged catchments, indirect characterization of hydrologic similarity becomes essential [*Blöschl*, 2006]. Two similarity approaches, viz., spatial proximity and physical similarity, have been shown to work successfully in many regions. In the spatial proximity approach, a gauged catchment that is located closest to the ungauged catchment is assumed to be hydrologically similar [*Mosley*, 1981; *Vandewiele et al.*, 1991; *Vandewiele and Elias*, 1995; *Merz and Blöschl*, 2004]; whereas in the physical similarity approach, a gauged catchment that is most similar to the ungauged catchment in physical attribute domain is assumed to be hydrologically similar [*Burn and Boorman*, 1993; *Parajka et al.*, 2005; *Oudin et al.*, 2010]. Studies that have compared these two approaches show that none has a clear advantage over the other for predicting streamflow at an ungauged catchment. For example, *Parajka et al.* [2005] used HBV model at 320 catchments in Austria and found that the physical similarity approach slightly outperformed the spatial proximity approach for catchments in Austria. On the other hand, *Oudin et al.* [2008] and *Zhang and Chiew* [2009] found that the spatial proximity approach performed marginally better than the physical similarity approach for estimating model parameters at ungauged catchments in France (913 catchments) and Australia (210 catchments), respectively.

Irrespective of the approach used, similarity-based procedures for parameter estimation typically involve transfer of all calibrated parameters from gauged to ungauged catchments [*Merz and Blöschl*, 2004; *McIntyre et al.*, 2005; *Oudin et al.*, 2008]. However, studies have shown that the identifiability of an optimal parameter value is not similar for all model parameters [*Beven*, 1989; *Beven and Binley*, 1992; *Doherty and Hunt*, 2009]. Specifically, some model parameters tend to be more sensitive to the objective function (i.e., their optimal values can be better constrained), whereas others can be insensitive over their entire feasible range. As a result, it is not entirely clear if all model parameters are equally important for transfer to ungauged catchments or if some parameters provide more hydrologically meaningful information than others, and should be preferentially transferred.

In this study, we hypothesize that there is some core information, contained within a subset of all the calibrated model parameters, whose transfer from gauged to ungauged catchments is the most critical factor for successful streamflow predictions. To test this hypothesis, we develop a simple daily time-step rainfall-runoff model (EXP-HYDRO) and implement it at 756 catchments across the continental United States. The EXP-HYDRO model contains six free calibration parameters. We first determine which model parameters are to be considered as important (or not important) based on their sensitivity to our objective function. We then compare the model performance at ungauged catchments by selectively transferring the different types of parameters. Both spatial proximity and physical similarity approaches are used to identify the donor gauged catchments for parameter transfer.

2. Data and Model

2.1 Data

We use daily streamflow data of 756 catchments from U. S. Geological Survey's Hydro-Climate Data Network (HCDN) (*Slack et al.*, [1993]; see Figure 1). The HCDN database consists of data of 1659 catchments located within the United States that are not severely affected by human activity and its record spans from 1874 to 1988. A majority of the catchments have consistent and continuous records from water year 1970 onwards. We consider only those catchments that have a continuous daily streamflow record from water year 1970 to 1988 (i.e., 1st October, 1969 to 30th September, 1988), which reduces the number of acceptable catchments to 756. The drainage area of the catchments ranges from 23 km^2 to 5100 km^2, whereas the average annual precipitation at the catchments ranges from 320 mm to 3300 mm.

Historical daily air temperature and precipitation data are obtained from the dataset developed by *Maurer et al.* [2002]. This data is gridded at 1/8 degree (about 14 km) spatial resolution and covers the entire continental United States. For each catchment, we also obtain data for five physio-climatic attributes from the dataset developed by *Vogel and Sankarasubramanian* [2005]. These attributes are: mean elevation above sea level, channel slope, soil permeability, solar radiation, percentage precipitation as snow. Aridity index (PET/P) is also calculated for each catchment from the available hydro-climatic data. Table 1 summarizes the physio-climatic attributes of all 756 catchments.

2.2 Rainfall-runoff Model

We have developed a simple spatially lumped rainfall-runoff model called EXP-HYDRO (EXPonential bucket HYDROlogic model). This model operates at a daily time-step and conceptualizes the catchment as a bucket store (Figure 2). The water balance equation of the catchment bucket is as follows:

$$\frac{dS}{dt} = P_r + M - ET - Q_{bucket} - Q_{spill} \tag{1}$$

where, S is the water stored in catchment bucket (unit: mm), P_r is the precipitation that falls as liquid rainfall (unit: mm/day), M is the snowmelt that occurs from the snow accumulation store (unit: mm/day). The snowmelt is modeled using a simple thermal degree-day model whose details are provided in Appendix A. ET is the evapotranspiration (unit: mm/day), Q_{bucket} is the runoff generated based on the available stored water in the bucket (unit: mm/day). Q_{spill} is the capacity excess runoff (unit: mm/day) that occurs only when excess precipitation and/or snowmelt is available to infiltrate into the catchment bucket, but the storage S has reached full capacity $S_{\max}$. The daily streamflow at catchment outlet is the sum of Q_{bucket} and Q_{spill}.

Evapotranspiration is calculated as a fraction of the potential evapotranspiration (*PET*), and depends on the amount of actual stored water relative to the bucket storage capacity:

$$ET = PET \cdot (S / S_{\max}) \tag{2}$$

PET (unit: mm/day) is obtained from daily air temperature using Hamon's formulation [*Hamon*, 1963]:

$$PET = 29.8 \cdot D \cdot \frac{e_{sat}(T_a)}{T_a + 273.2} \tag{3}$$

where, D is the day length (unit: hours), which depends on the Julian date of the year and the latitude of catchment location. D is calculated using the formula suggested by *Dingman* [2002] (Appendix E in that book). e_{sat} is the saturation vapor pressure (unit: kPa), calculated as:

$$e_{sat}(T_a) = 0.611 \cdot \exp\left(\frac{17.3 \cdot T_a}{T_a + 237.3} \right) \tag{4}$$

The amount of runoff generated from the catchment bucket depends on the amount of water stored in it and is calculated using a TOPMODEL [*Beven and Kirkby*, 1979] type equation:

$$Q_{bucket} = Q_{\max} \cdot \exp(-f \cdot (S_{\max} - S)) \tag{5}$$

where, $Q_{\max}$ is the maximum runoff produced by the catchment bucket (unit: mm/day) when its storage reaches the maximum capacity, and f is the parameter controlling the storage-dependent decline in runoff (unit: mm^{-1}).

3. Methods

In this section, we first outline the procedure used for calibration of EXP-HYDRO model parameters at the 756 catchments. Then we describe the method used for identifying and classifying the sensitive and insensitive model parameters. This is followed by a brief description of the parameter transfer schemes used for estimating model parameters at ungauged catchments.

3.1 Model Calibration

The EXP-HYDRO model contains six free calibration parameters (f, $Q_{\max}$, $S_{\max}$, D_f, $T_{\min}$, and $T_{\max}$), of which, D_f, $T_{\min}$, and $T_{\max}$ are the parameters from snow model (see Appendix A). For each catchment, we calibrate the above six parameters with 50,000 Monte Carlo simulations. Table 2 shows the parameter ranges used for random sampling of these six parameters. First year from the chosen time-period (water year 1970) is used for model spin-up, and the daily streamflow data from remaining 18 years is used for parameter optimization. Nash Sutcliffe efficiency [*Nash and Sutcliffe*, 1970] of square root values of daily streamflow is used as the objective function:

$$NS = 1 - \frac{\sum_{i=1}^{n} (\sqrt{Q_{obs,i}} - \sqrt{Q_{pred,i}})^2}{\sum_{i=1}^{n} (\sqrt{Q_{obs,i}} - \sqrt{\overline{Q}_{obs}})^2} \qquad (6)$$

where, $Q_{pred,i}$ and $Q_{obs,i}$ are the predicted and the observed streamflow values on the i^{th} day respectively, $\overline{Q}_{obs}$ is the mean of all the observed streamflow values and n is the total number of days in the record. The commonly used variants of Nash-Sutcliffe efficiency formula are: untransformed (Q), square root transformed ($\sqrt{Q}$), and log transformed ($log\ Q$) [*Oudin et al.*, 2006]. As an objective function, NS (Q) tends to over-emphasize the matching of high flow values (at the expense of low flows), whereas NS ($log\ Q$) tends to do the opposite. NS ($\sqrt{Q}$), however, balances out these two extremes and focuses on matching the overall hydrograph, albeit at the expense of very high and very low flows. Since our objective in this study is to match the overall hydrologic dynamics of a catchment over a long time period, we use NS ($\sqrt{Q}$) as the objective function (Equation 6, and henceforth referred to simply as NS). Comparison of different objective functions is beyond the scope of this study.

3.2 Model parameter sensitivity

To characterize the sensitivity of EXP-HYDRO model parameters, we implement a simple procedure that tests the improvement in model performance when a given parameter is assigned its calibrated value instead of a randomly sampled value. We begin with a baseline scenario where the values of all six parameters are randomly sampled within their feasible ranges (see Table 2). This baseline scenario is illustrated as Run 1 in Figure 3, where the solid gray bar shows the median NS of all 756 catchments and the error bars show the 25th and 75th percentile values of median NS (obtained through 1000 iterative model runs). We next fix each model parameter individually to its calibrated value (while still keeping the other five parameters random) and measure the increase in model performance from the baseline scenario. The maximum increase in median NS is obtained when f is fixed to its calibrated value (Run 2 in Figure 3), which indicates that f is the most sensitive parameter in the EXP-HYDRO model. For Run 3, we keep f fixed and repeat the procedure from Run 2 by individually fixing each of the remaining five parameters to calibrated values. The largest increase in median NS during Run 3 is obtained when both f and S_{max} are fixed (see Figure 3). This suggests that S_{max} is the second most sensitive parameter in the EXP-HYDRO model. The next largest increase in median NS is obtained when f, S_{max}, and D_f are fixed to their calibrated values (Run 4 in Figure 3). To identify the fourth most sensitive parameter, we next add the calibrated values of Q_{max}, T_{min}, and T_{max}

individually to the list of fixed parameters (Runs 5, 6, and 7, respectively). However, we observe that the increase in median NS is much smaller, and similar, when either of these parameters is fixed to the calibrated values (Figure 3). This suggests that these three parameters are equally sensitive (or insensitive) to the objective function. Therefore, we classify parameters f, S_{max}, and D_f as sensitive parameters, and Q_{max}, T_{min}, and T_{max} as insensitive parameters.

The above mentioned classification of parameters is consistent with visual observation of the dotty plots (widely used in GLUE methodology [*Beven and Binley*, 1992]) from the 50,000 Monte Carlo simulations used for calibration. Figure 4 shows the dotty plots of all six parameters for two contrasting catchments; rain dominated (in western Oregon) and snow dominated (in Wyoming). It can be noted from this figure that the dot density is much higher for the Oregon catchment, which suggests that it has a larger number of parameter combinations that yield high NS values. In both catchments, f is the most sensitive parameter with a narrow range of values that produce high NS. While S_{max} is less sensitive to NS than f, certain value ranges of the S_{max} parameter appear to be unfavorable for obtaining high NS. For the snow-dominated catchment in Wyoming, we find that, in addition to f and S_{max}, parameter D_f from the snow model shows sensitivity to NS (Figure 4b). All three parameters from the snow component are insensitive to NS at the rain-dominated catchment in Oregon (Figure 4a). This is expected since the snow component of the EXP-HYDRO model will mostly be inactive in this catchment. However, from the three parameters of the catchment bucket, we find that only Q_{max} remains insensitive over its entire range in both catchments.

3.3 Parameter estimation at ungauged catchments

We use both physical similarity and spatial proximity based approaches to identify the donor gauged catchments for parameter transfer. In the physical similarity approach, physio-climatic attributes of each catchment are obtained, and the catchment that is most similar to the ungauged catchment in physical attribute domain is chosen as the donor catchment for parameter transfer. We consider seven catchment attributes: drainage area, mean elevation, channel slope, soil permeability, solar radiation, percentage precipitation as snow, and aridity index (P/PET). The attribute distance between the catchments is calculated as follows:

$$dist_{a,b} = \sqrt{\sum_{j=1,J}\left(\frac{X_{a,j} - X_{b,j}}{\max(X_j) - \min(X_j)}\right)^2} \quad (7)$$

where, J is the total number of catchment attributes (J = 7 in our case), $X_{a,j}$ is the value of an attribute at catchment a, and $\max(X_j) - \min(X_j)$ is the range of that attribute among all the catchments considered. Gauged catchment with the lowest value of *dist* is chosen as the donor catchment. In the spatial proximity approach, only geographic distance among catchments is considered. We use the Euclidean distance between the stream gauge locations to quantify spatial proximity. Gauged catchment that is located closest to the ungauged catchment is chosen as the donor catchment.

Using the above two approaches, we test four different schemes of parameter estimation at an ungauged catchment. In scheme 1, all six parameters of the EXP-HYDRO model are transferred from the donor gauged catchment. In scheme 2, only the sensitive parameters (f, $S_{\max}$, and D_f) are transferred from the gauged catchment and the insensitive parameters ($Q_{\max}$, $T_{\min}$, and $T_{\max}$) are assigned a random value within their parameter range (see Table 2). In scheme 3, the sensitive parameters are chosen randomly within their parameter range and the insensitive parameters are transferred from gauged catchment. In scheme 4, all six parameters are chosen randomly within their parameter range and no information is transferred from the donor gauged catchment to an ungauged catchment.

4. Results

4.1 Model performance at gauged catchments

Based on the calibration performance of EXP-HYDRO model, we first identify the gauged catchments that meet our criterion for acceptable performance (NS > 0.6). If a catchment can be calibrated with NS > 0.6, we consider the structure of EXP-HYDRO model to be suitable for the simulation of hydrologic dynamics at that catchment. While this criterion is subjective in nature, our observations of simulated hydrographs at numerous catchments suggested that, even if each event is not simulated accurately, hydrographs with NS > 0.6 can reliably mimic the overall observed hydrologic patterns across several years. Figure 5 shows the geographic distribution of catchments that are "accepted" and "rejected" based on our criterion. Out of the 756 catchments, 323 catchments (~ 43%) have calibrated model performance with NS > 0.6. Majority of the accepted catchments are located in three distinct geographic regions: (1) in the eastern half of the US, mainly along Appalachian Mountain Range, but also in some mid-western and southern states on either side of the Mississippi river; (2) along the Rocky Mountain Range in the states of Idaho, Wyoming, and Colorado; and (3) along the Pacific coast, to the west of the Cascade and the Sierra Nevada Mountain Ranges. On the other hand, majority of the rejected catchments (NS < 0.6) are located in the drier central part of the US, the rain-shadow regions in western US, along the Gulf Coast, and to the east of the Appalachian Mountain Range in Mid-Atlantic States.

4.2 Model performance at ungauged catchments

For the transfer of model parameters to the ungauged catchments, we only consider the 323 accepted catchments where we know a priori that the EXP-HYDRO model structure is suitable. Each of the 323 catchments is considered ungauged in turn, its donor gauged catchment is chosen (based on either spatial proximity or physical similarity), and appropriate model information is transferred to this pseudo-ungauged catchment based on the four parameter estimation schemes described in Section 3.3.

We first compare the spatial proximity and physical similarity approaches in terms of model performance at ungauged catchments. For direct comparison of these two approaches, we only consider model predictions from scheme 1, where all six parameters are transferred from gauged to ungauged catchments. Figure 6a shows the empirical CDF (cumulative distribution function) plot of the NS values for calibration case (blue line), spatial proximity based parameter transfer (red line), and physical similarity based parameter transfer (black line). Both spatial proximity and physical similarity approaches provide similar overall model performance, but NS values are slightly higher for the spatial proximity approach at high percentiles. Figure 6b provides a 1:1 comparison of the NS values from these two parameter transfer approaches. In 202 catchments (out of 323 in total; ~ 63%), NS values are equal or higher with the spatial proximity approach than with the physical similarity approach. Figure 7 shows the map of catchments where either of these two approaches performs better. We find no distinct geographic regions where one approach has a complete advantage over the other. In terms of acceptable model performance, 187 catchments (~ 58%) have NS > 0.6 using the spatial proximity approach, whereas 179 catchments (~ 55%) have NS > 0.6 using the physical similarity approach.

Next, we compare the model performance at ungauged catchments using the four parameter estimation schemes (see Section 3.3). Our goal in testing these schemes is to determine if transfer of sensitive parameters from gauged catchments is more valuable for streamflow prediction at ungauged catchments than the transfer of insensitive parameters. Figure 8 compares these four schemes through the empirical CDF plots and box plots of NS values. For both spatial proximity and physical similarity approaches, we find that although scheme 1 (all six parameters transferred) provides the best overall predictability for the pseudo-ungauged scenario, scheme 2 (sensitive parameters transferred; insensitive parameters chosen randomly within parameter range) provides predictability that is almost on par with scheme 1, especially at higher percentiles of NS. Scheme 3 (insensitive parameters transferred; sensitive parameters chosen randomly within parameter range) and scheme 4 (all six parameters chosen randomly within parameter range) provide a model performance that is significantly deteriorated compared to the performance from schemes 1 and 2. It is worth noting here that when only insensitive parameters are transferred, the model performance is almost equivalent to that of using a completely randomized parameter set. Box plot comparison (Figure 8) of these four schemes shows that the variability of NS among the 323 catchments is substantially smaller (and almost similar) for schemes 1 and 2, compared to that for schemes 3 and 4.

5. Discussion

In terms of parameter transfer from gauged to ungauged catchments, regionalization studies in the past have not treated sensitive and insensitive model parameters differently. Some have even recommended transfer of the entire calibrated parameter set to ensure that internal dependencies or correlations among optimized model parameters are preserved [*McIntyre et al.*, 2005; *Oudin et al.*, 2008]. *Kokkonen et al.* [2003] stated that "…when there is a reason to believe that, in the sense of hydrological behaviour, a gauged catchment resembles the ungauged catchment to a sufficient extent, it may be worthwhile to adopt the entire set of calibrated parameters from the gauged catchment". While our results are in general agreement with this recommendation, they certainly reveal that major differences exist when different type of model information is transferred selectively. Specifically, we find that the success of streamflow prediction at an ungauged catchment depends largely on the transfer of model parameters that are sensitive to our objective function (NS). On the other hand, transfer of insensitive model parameters from the donor gauged catchments is significantly less valuable if sensitive parameters are not well estimated in the first place. The importance of sensitive parameters at gauged catchments is obvious, since deviations from the optimal values will likely result in significant performance decline (see Figure 3). It is less intuitive, however, that these same (sensitive) parameters would retain their importance when transferring information from gauged to ungauged catchments. This suggests that the sensitive model parameters not only contain information that controls the hydrologic behavior at gauged catchments, but they also determine the extent to which streamflow predictability can be achieved at an ungauged catchment in the region.

Although identifying the exact hydrologic information contained in calibrated model parameters can be difficult in some cases, the individual role played by each parameter within the model structure offers clues into the hydrologic processes that they represent. All the three sensitive parameters of EXP-HYDRO model convey different aspects of the hydrograph recession information. Parameters f and D_f essentially control the rate of depletion of water and snow storage reservoirs within the catchment. $S_{\max}$, on the other hand, represents an effective depth within the catchment at which flow contribution to the stream ceases. Small value of $S_{\max}$ indicates a shallow system that is most likely dominated by flow paths with short residence times, whereas large value of $S_{\max}$ suggests a deep system that allows for greater contribution from slower flow paths. This phenomenon is noticeable in Figure 9 which shows that an inverse relationship exists between f and S_{max} for the 323 accepted catchments. Specifically, the rate of depletion f tends to be slower for a deep bucket (high S_{max}), which prolongs the hydrograph recession due to greater contributions from slower (and perhaps deeper) flow paths. A shallow bucket (low S_{max}) tends to show the opposite behavior where a quicker depletion of the hydrograph recession limb occurs. The presence of insensitive parameters in a model might reflect an inadequate understanding or representation of some hydrologic processes within the model structure. For instance, Q_{max} represents the maximum flow contribution from the catchment bucket when it is completely saturated. Ideally, this parameter can be well constrained since it is conceptually related to the lateral conductivity of a saturated soil column. However, due to our incomplete knowledge of the internal heterogeneity and macropore structure of soils within the catchment, this parameter might have become insensitive in practice. Parameters T_{max} and T_{min} also show insensitivity to the objective function. A likely reason for this could be the simplistic representation of snow processes in the thermal degree-day snow model, which is solely dependent on the surface air temperature. A more complex representation of the snow accumulation and melt processes might help in reducing the insensitivity of such snow-related parameters.

Comparison of the spatial proximity and physical similarity approaches showed that almost similar model performance is achieved with either approach. This is consistent with previous studies that have compared these two approaches, but by using different combinations of physio-climatic attributes for the physical similarity approach [*Parajka et al.*, 2005; *Oudin et al.*, 2008; *Zhang and Chiew*, 2009]. For instance, *Oudin et al.* [2008] used six attributes, viz., catchment area, catchment slope, median altitude, drainage density, fraction of forest cover, and aridity index. *Zhang and Chiew* [2009] used eight attributes, such as area, aridity index, mean elevation, mean slope, stream length, mean solum thickness, plant available water holding capacity, and mean woody vegetation fraction. These differences are reflective of the disparity that exists in available geophysical data from various parts of the world. Regardless of the combination used, however, a physical similarity based framework typically contains both physiographic and climatic attributes. To gain further insight into the relative influence of each, we compare the model performance when only physiographic vs. climatic attributes are used to identify donor catchments. Figure 10a shows the CDF plot of NS values for the 323 pseudo-ungauged catchments with three scenarios for selecting a donor gauged catchment: (1) all seven attributes are used, (2) only climatic attributes (aridity index, solar radiation, percentage precipitation as snow) are used, and (3) only physiographic attributes (drainage area, channel slope, mean elevation, soil permeability) are used. We find that the model performance with considering climatic attributes only is marginally better than that with considering physiographic attributes only. This suggests that, on their own, the climatic attributes have slightly more explanatory power regarding catchment similarity than our chosen physiographic attributes. One reason could be that a stronger connection exists between climatic similarity and spatial proximity, i.e., catchments having similar climate are more likely to be located close to each other. Figure 10b shows a 1:1 comparison of the NS values obtained with climatic and physiographic attributes. While most catchments have NS values close to the 1:1 line, large scatter in this relationship suggests that climatic similarity is clearly preferred to physiographic similarity (and vice versa) in some catchments. Nonetheless, the best performance is still achieved when both physiographic and climatic attributes are used within a catchment similarity framework.

The EXP-HYDRO model developed in this study performs satisfactorily (NS > 0.6) in only 43% of the 756 catchments. Nonetheless, the geographic distribution of good predictability catchments (Figure 5) is similar to that observed by previous modeling studies within the continental US, even though completely different models and temporal resolution (monthly) were used in these studies [*Hay and McCabe*, 2002; *Martinez and Gupta*, 2010]. We think that any other model which is implementable across a large number of catchments will likely produce similar geographic patterns of streamflow predictability. The method that we used to identify important vs. non-important information within the EXP-HYDRO model is based on our observation of parameter sensitivity to a single objective function (NS). It is likely that the sensitivity of a model parameter will be different if other objective functions are used, in which case a completely different set of parameters will become important. Regardless of the objective function used, however, a modeler will have to analyze the role played by a particular parameter in representing the function of the system, and then take a decision as to whether that parameter conveys meaningful information or not. Overall, we think that the findings from this study are generic enough in nature and applicable to any modeling framework. While parameters with different sensitivities will almost always exist in any model structure, identifying the key information that controls model behavior will certainly lead to progress in our understanding of the hydrologic systems.

6. Concluding remarks

In this study, we tested the hypothesis that there is some core information, contained within a subset of all calibrated model parameters, whose transfer from gauged to ungauged catchments is the most critical factor for successful streamflow predictions. To this end, we developed a simple daily time-step rainfall-runoff model (EXP-HYDRO) and implemented it over 756 catchments across the continental United States. Both spatial proximity and physical similarity based approaches were tested to identify the donor gauged catchments for parameter transfer. Based on the results, we conclude that streamflow predictability at ungauged catchments using rainfall-runoff models is largely dependent on the transfer of a small subset of parameters from donor gauged catchments. In the case of EXP-HYDRO model, this subset consists of three parameters that convey different aspects of the hydrograph recession information, and are also sensitive to our objective function (NS). Importantly, we find that the transfer of this key information is essential regardless of the approach used for identifying the donor gauged catchments. We further recommend that, in any modeling framework, the core subset of important parameters should be identified and better characterized in order to understand the information that controls predictability at ungauged catchments.

Acknowledgements

This research was supported by the following NSF grants: ARC-0922100 and EAR-1027870. The authors would like to thank Dr. Malcolm Anderson (editor) for facilitating a thorough review, and the anonymous reviewers for providing insightful comments that have greatly improved this manuscript.

Appendix A: Thermal degree-day snow model

The EXP-HYDRO model contains two buckets, a catchment bucket and a snow accumulation bucket. Only the precipitation that is considered as snowfall accumulates in the snow bucket, whereas the rainfall accumulates directly in the catchment bucket. The daily precipitation P is classified as snowfall or rainfall based on the following conditions:

If $T_a < T_{\min}$,

$$P_s = P \\ P_r = 0 \qquad \text{(A1a)}$$

Else,

$$P_s = 0 \\ P_r = P \qquad \text{(A1b)}$$

where, P_s is snowfall in mm/day, P_r is rainfall in mm/day, and T_a is daily air temperature in °C. Water balance of the snow bucket is as follows:

$$\frac{dS_{snow}}{dt} = P_s - M \qquad \text{(A2)}$$

Where, S_{snow} is the storage in snow bucket (unit: mm), and M is the snowmelt (unit: mm/day). The amount of snowmelt M is modeled using the thermal degree-day concept as follows:

If $S_{snow} > 0$ and $T_a > T_{\max}$,

$$M = \min\{S_{snow}, D_f \cdot (T_a - T_{\max})\} \qquad \text{(A3a)}$$

Else,

$$M = 0 \qquad \text{(A3b)}$$

where, D_f is the thermal degree-day factor (unit: mm/day/°C), and $T_{\max}$ is the temperature threshold above which accumulated snow begins to melt. The snowmelt M from the snow bucket is input to the catchment bucket (see Equation 1).

References

Abdelnour, A., M. Stieglitz, F. Pan, and R. McKane (2011), Catchment hydrological responses to forest harvest amount and spatial pattern, *Water Resources Research*, *47*(9), W09521, doi: 10.1029/2010wr010165.

Archfield, S. A., and R. M. Vogel (2010), Map correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments, *Water Resources Research*, *46*(10), W10513, doi: 10.1029/2009wr008481.

Beven, K. J. (1989), Changing ideas in hydrology — The case of physically-based models, *Journal of Hydrology*, *105*(1–2), 157-172, doi: 10.1016/0022-1694(89)90101-7.

Beven, K. J. (2001), *Rainfall-Runoff Modelling: The Primer*, 319 pp., John Wiley & Sons, Ltd., Chichester.

Beven, K. J., and M. J. Kirkby (1979), A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant, *Hydrological Sciences Bulletin*, *24*(1), 43-69, doi: 10.1080/02626667909491834.

Beven, K. J., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrological Processes*, *6*(3), 279-298, doi: 10.1002/hyp.3360060305.

Blöschl, G. (2006), Rainfall-Runoff Modeling of Ungauged Catchments, in *Encyclopedia of Hydrological Sciences*, edited, John Wiley & Sons, Ltd, doi: 10.1002/0470848944.hsa140.

Burn, D. H., and D. B. Boorman (1993), Estimation of hydrological parameters at ungauged catchments, *Journal of Hydrology*, *143*(3-4), 429-454, doi: 10.1016/0022-1694(93)90203-l.

Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer (2004), The Effects of Climate Change on the Hydrology and Water Resources of the Colorado River Basin, *Climatic Change*, *62*(1), 337-363, doi: 10.1023/B:CLIM.0000013684.13621.1f.

Dingman, S. L. (2002), *Physical hydrology*, 600 pp., Prentice Hall, New Jersey.

Doherty, J., and R. J. Hunt (2009), Two statistics for evaluating parameter identifiability and error reduction, *Journal of Hydrology*, *366*(1–4), 119-127, doi: 10.1016/j.jhydrol.2008.12.018.

Hamlet, A. F., and D. P. Lettenmaier (1999), Effects of Climate Change on Hydrology and Water Resources in the Columbia River Basin, *Journal of the American Water Resources Association*, *35*(6), 1597-1623, doi: 10.1111/j.1752-1688.1999.tb04240.x.

Hamon, W. R. (1963), Computation of direct runoff amounts from storm rainfall, *Int. Assoc. Sci. Hydrol. Publ*, *63*, 52–62, doi.

Hay, L. E., and G. J. McCabe (2002), Spatial variability in water-balance model performance in the conterminous United States, *Journal of the American Water Resources Association*, *38*(3), 847-860, doi: 10.1111/j.1752-1688.2002.tb01001.x.

Kokkonen, T. S., A. J. Jakeman, P. C. Young, and H. J. Koivusalo (2003), Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina, *Hydrological Processes*, *17*(11), 2219-2238, doi: 10.1002/hyp.1329.

Krysanova, V., D.-I. Müller-Wohlfeil, and A. Becker (1998), Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds, *Ecological Modelling*, *106*(2-3), 261-289, doi: 10.1016/s0304-3800(97)00204-4.

Martinez, G. F., and H. V. Gupta (2010), Toward improved identification of hydrological models: A diagnostic evaluation of the "abcd" monthly water balance model for the conterminous United States, *Water Resources Research*, *46*(8), W08507, doi: 10.1029/2009wr008294.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States, *Journal of Climate*, *15*(22), 3237-3251, doi: 10.1175/1520-0442(2002)015<3237:althbd>2.0.co;2.

McIntyre, N., H. Lee, H. Wheater, A. Young, and T. Wagener (2005), Ensemble predictions of runoff in ungauged catchments, *Water Resources Research*, *41*(12), W12434, doi: 10.1029/2005wr004289.

Merz, R., and G. Blöschl (2004), Regionalisation of catchment model parameters, *Journal of Hydrology*, *287*(1-4), 95-123, doi: 10.1016/j.jhydrol.2003.09.028.

Merz, R., and G. Blöschl (2005), Flood frequency regionalisation—spatial proximity vs. catchment attributes, *Journal of Hydrology*, *302*(1-4), 283-306, doi: 10.1016/j.jhydrol.2004.07.018.

Moretti, G., and A. Montanari (2008), Inferring the flood frequency distribution for an ungauged basin using a spatially distributed rainfall-runoff model, *Hydrology and Earth System Sciences*, *12*(4), 1141-1152, doi: 10.5194/hess-12-1141-2008.

Mosley, M. P. (1981), Delimitation of New Zealand hydrologic regions, *Journal of Hydrology*, *49*(1-2), 173-192, doi: 10.1016/0022-1694(81)90211-0.

Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I — A discussion of principles, *Journal of Hydrology*, *10*(3), 282-290, doi: 10.1016/0022-1694(70)90255-6.

Notter, B., H. Hurni, U. Wiesmann, and K. C. Abbaspour (2012), Modelling water provision as an ecosystem service in a large East African river basin, *Hydrology and Earth System Sciences*, *16*(1), 69-86, doi: 10.5194/hess-16-69-2012.

Oudin, L., A. Kay, V. Andréassian, and C. Perrin (2010), Are seemingly physically similar catchments truly hydrologically similar?, *Water Resources Research*, *46*(11), W11558, doi: 10.1029/2009wr008887.

Oudin, L., V. Andréassian, T. Mathevet, C. Perrin, and C. Michel (2006), Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations, *Water Resources Research*, *42*(7), W07410, doi: 10.1029/2005wr004636.

Oudin, L., V. Andréassian, C. Perrin, C. Michel, and N. Le Moine (2008), Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments, *Water Resources Research*, *44*(3), W03413, doi: 10.1029/2007wr006240.

Parajka, J., R. Merz, and G. Blöschl (2005), A comparison of regionalisation methods for catchment model parameters, *Hydrology and Earth System Sciences*, *9*(3), 157-171, doi: 10.5194/hess-9-157-2005.

Patil, S., and M. Stieglitz (2012), Controls on hydrologic similarity: role of nearby gauged catchments for prediction at an ungauged catchment, *Hydrology and Earth System Sciences*, *16*(2), 551-562, doi: 10.5194/hess-16-551-2012.

Poff, N. L., B. D. Richter, A. H. Arthington, S. E. Bunn, R. J. Naiman, E. Kendy, M. Acreman, C. Apse, B. P. Bledsoe, M. C. Freeman, J. Henriksen, R. B. Jacobson, J. G. Kennen, D. M. Merritt, J. H. O'Keeffe, J. D. Olden, K. Rogers, R. E. Tharme, and A. Warner (2010), The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards, *Freshwater Biology*, *55*(1), 147-170, doi: 10.1111/j.1365-2427.2009.02204.x.

Rees, H. G., M. G. R. Holmes, A. R. Young, and S. R. Kansakar (2004), Recession-based hydrological models for estimating low flows in ungauged catchments in the Himalayas, *Hydrology and Earth System Sciences*, *8*(5), 891-902, doi: 10.5194/hess-8-891-2004.

Servais, P., G. Billen, A. Goncalves, and T. Garcia-Armisen (2007), Modelling microbiological water quality in the Seine river drainage network: past, present and future situations, *Hydrology and Earth System Sciences*, *11*(5), 1581-1592, doi: 10.5194/hess-11-1581-2007.

Shukla, S., and D. P. Lettenmaier (2011), Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill, *Hydrology and Earth System Sciences*, *15*(11), 3529-3538, doi: 10.5194/hess-15-3529-2011.

Singh, V. P. (1995), *Computer models of watershed hydrology*, 1144 pp., Water Resources Publications, Highlands Ranch, CO.

Singh, V. P., and D. A. Woolhiser (2002), Mathematical Modeling of Watershed Hydrology, *Journal of Hydrologic Engineering*, *7*(4), 270-292, doi: 10.1061/(ASCE)1084-0699(2002)7:4(270).

Singh, V. P., and D. K. Frevert (2006), *Watershed Models*, Taylor & Francis, Boca Raton, FL.

Sivapalan, M., K. Takeuchi, S. W. Franks, V. K. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J. J. McDonnell, E. M. Mendiondo, P. E. O'Connell, T. Oki, J. W. Pomeroy, D. Schertzer, S. Uhlenbrook, and E. Zehe (2003), IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, *Hydrological Sciences Journal*, *48*(6), 857-880, doi: 10.1623/hysj.48.6.857.51421.

Slack, J. R., A. Lumb, and J. M. Landwehr (1993), Hydro-Climatic Data Network (HCDN) Streamflow Data Set, 1874-1988: *USGS Water-Resources Investigations Report 93-4076*, U.S. Geological Survey, Reston, VA.

Smakhtin, V. U. (2001), Low flow hydrology: a review, *Journal of Hydrology*, *240*(3-4), 147-186, doi: 10.1016/s0022-1694(00)00340-1.

Staudinger, M., K. Stahl, J. Seibert, M. P. Clark, and L. M. Tallaksen (2011), Comparison of hydrological model structures based on recession and low flow simulations, *Hydrology and Earth System Sciences*, *15*(11), 3447-3459, doi: 10.5194/hess-15-3447-2011.

Vandewiele, G. L., and A. Elias (1995), Monthly water balance of ungauged catchments obtained by geographical regionalization, *Journal of Hydrology*, *170*(1-4), 277-291, doi: 10.1016/0022-1694(95)02681-e.

Vandewiele, G. L., C.-Y. Xu, and W. Huybrechts (1991), Regionalisation of physically-based water balance models in Belgium. Application to ungauged catchments, *Water Resources Management*, *5*(3), 199-208, doi: 10.1007/bf00421989.

Vogel, R. M., and A. Sankarasubramanian (2005), Monthly Climate Data for Selected USGS HCDN Sites, 1951–1990, edited, Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.

Wagener, T., and A. Montanari (2011), Convergence of approaches toward reducing uncertainty in predictions in ungauged basins, *Water Resources Research*, *47*(6), W06301, doi: 10.1029/2010wr009469.

Xu, H., R. G. Taylor, and Y. Xu (2011), Quantifying uncertainty in the impacts of climate change on river discharge in sub-catchments of the Yangtze and Yellow River Basins, China, *Hydrology and Earth System Sciences*, *15*(1), 333-344, doi: 10.5194/hess-15-333-2011.

Zealand, C. M., D. H. Burn, and S. P. Simonovic (1999), Short term streamflow forecasting using artificial neural networks, *Journal of Hydrology*, *214*(1-4), 32-48, doi: 10.1016/s0022-1694(98)00242-x.

Zhang, Y., and F. H. S. Chiew (2009), Relative merits of different methods for runoff predictions in ungauged catchments, *Water Resources Research*, *45*(7), W07412, doi: 10.1029/2008wr007504.

Tables

Table 1: Distribution of the physio-climatic properties among 756 catchments

	Area (km^2)	Elevation* (m)	Channel Slope* (degrees)	Solar Radiation* (mm/yr)	PPS* (%)	Permeability* (mm)	Aridity Index
Max	5102.30	3646.40	13.57	4830.60	71.93	166.20	3.07
75^{th} %ile	1665.40	785.48	3.37	4467.80	12.53	41.54	0.99
50^{th} %ile	748.51	382.06	0.75	4246.20	6.02	25.25	0.78
25^{th} %ile	310.80	232.53	0.36	4073.40	1.21	15.68	0.65
Min	23.31	7.19	0.03	3700.20	0.00	4.68	0.24

* Data obtained from *Vogel and Sankarasubramanian* [2005] dataset

Table 2: Parameter ranges for calibration of EXP-HYDRO model

Parameter	Description	Units	Lower Limit	Upper Limit
f	Rate of decline in runoff from catchment bucket	mm^{-1}	0.0	0.1
S_{max}	Maximum storage of the catchment bucket	mm	100.0	1500.0
Q_{max}	Maximum subsurface runoff at full bucket	mm/day	10.0	50.0
D_f	Thermal degree-day factor	mm/day/°C	0.0	5.0
T_{max}	Temperature above which snow starts melting	°C	0.0	3.0
T_{min}	Temperature below which precipitation is snow	°C	-3.0	0.0

Figures

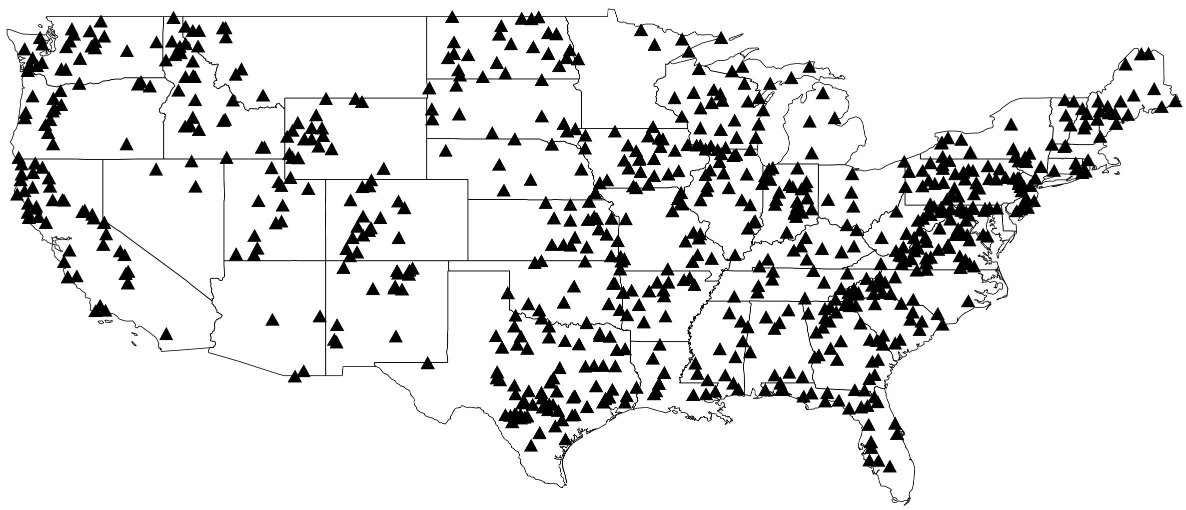

Figure 1: Location of the 756 study catchments within continental United States.

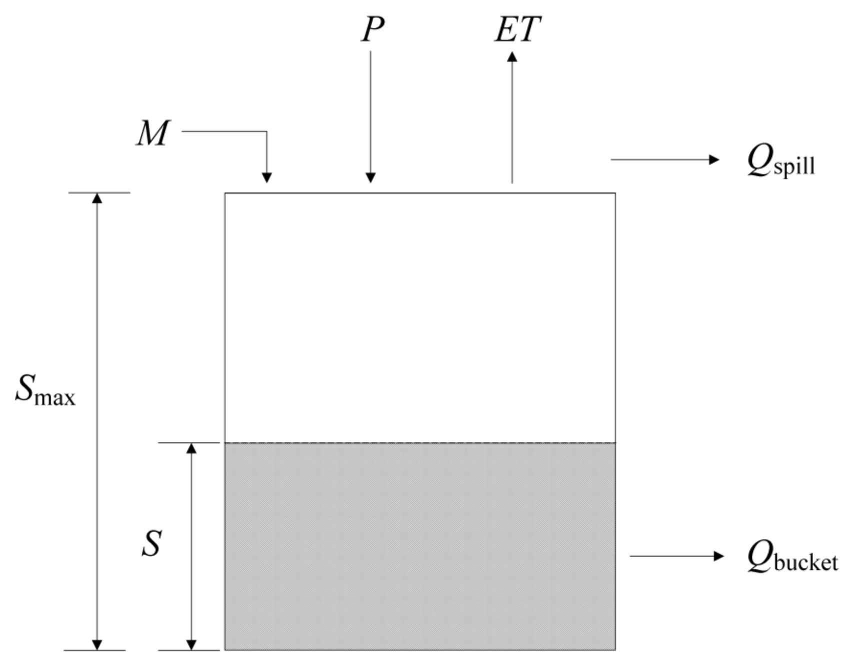

Figure 2: Schematic representation of the EXP-HYDRO rainfall-runoff model.

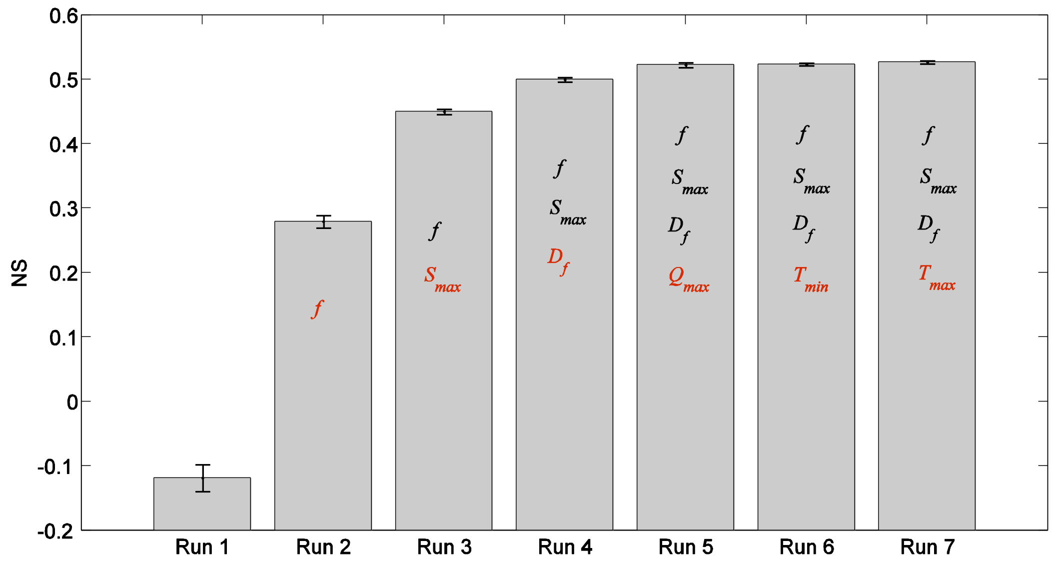

Figure 3: NS values obtained with multiple parameter estimation scenarios. Gray bar denotes median NS among the 756 catchments, and the error bars denote the 25th and 75th percentile values of median NS (obtained through 1000 iterative model runs). For model runs 2 to 7, parameter highlighted in red is the new parameter that is fixed to its calibrated value compared to previous model runs.

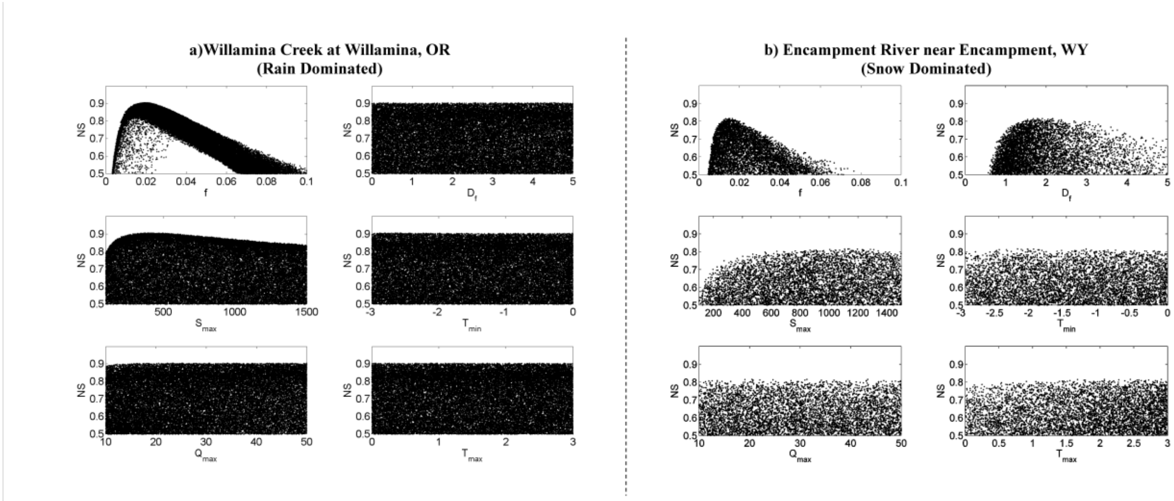

Figure 4: Dotty plots of the model parameters from 50,000 Monte Carlo simulations for a) Willamina Creek in Oregon (rain dominated), and b) Encampment River in Wyoming (snow dominated).

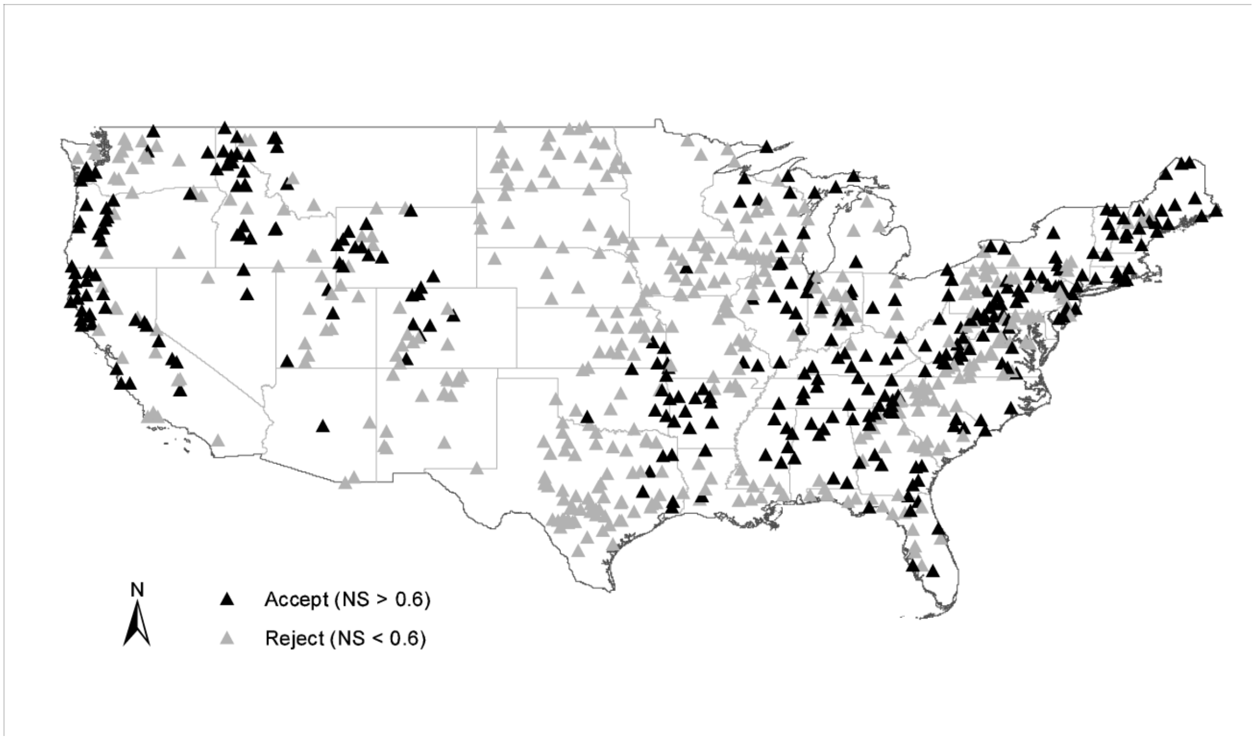

Figure 5: Location of the catchments that are either “accepted” or “rejected” based on the model performance criterion.

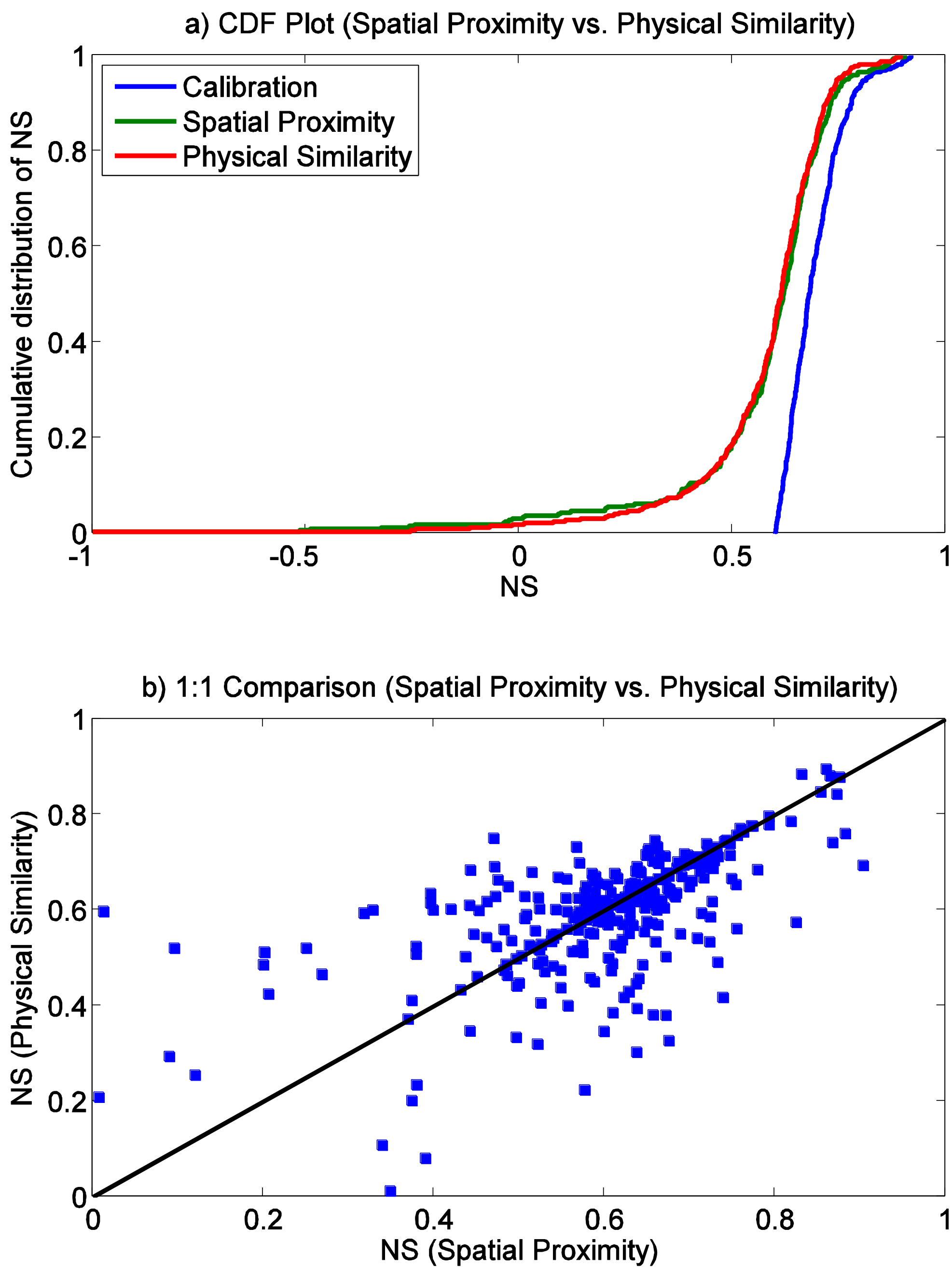

Figure 6: Comparison of the spatial proximity and physical similarity approaches with a) CDF plot, and b) 1:1 comparison of NS values.

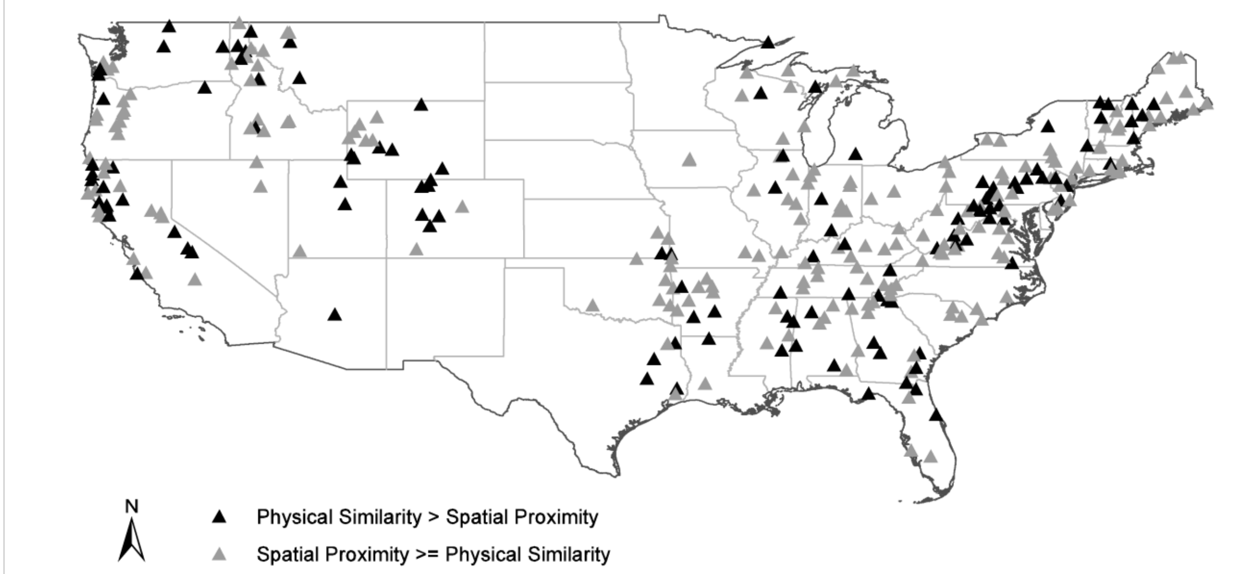

Figure 7: Location of the catchments where the model performance with spatial proximity approach is better (or equal) and worse than the physical similarity approach.

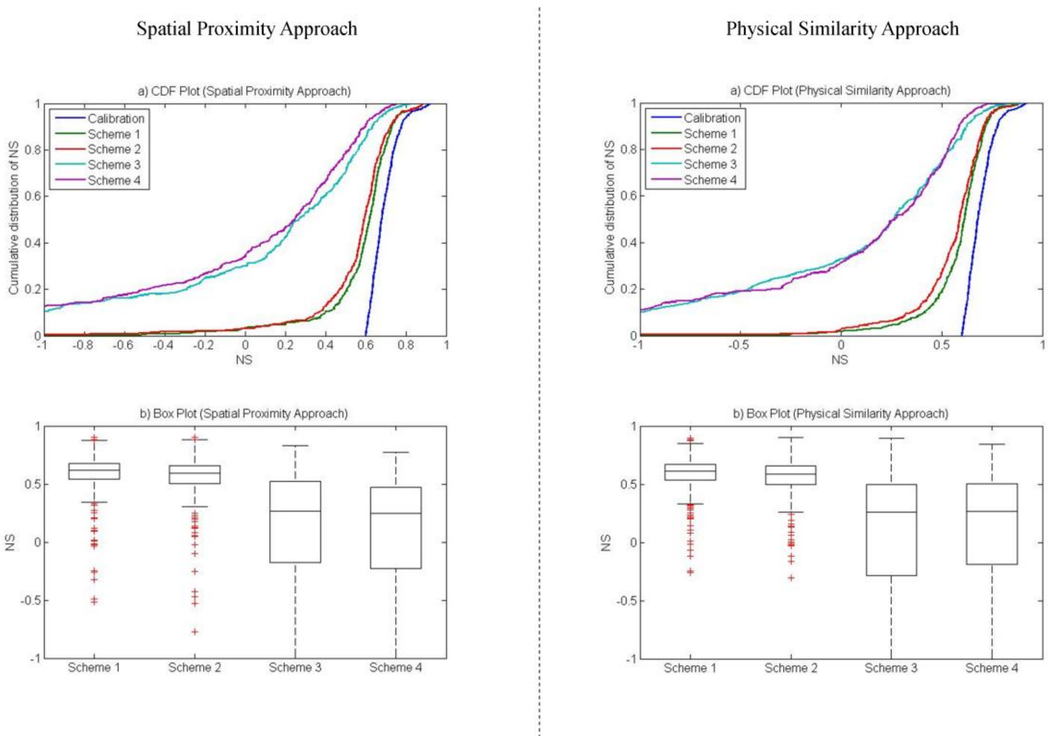

Figure 8: Comparison of model performance at ungauged catchments with the four parameter transfer schemes for spatial proximity and physical similarity approaches.

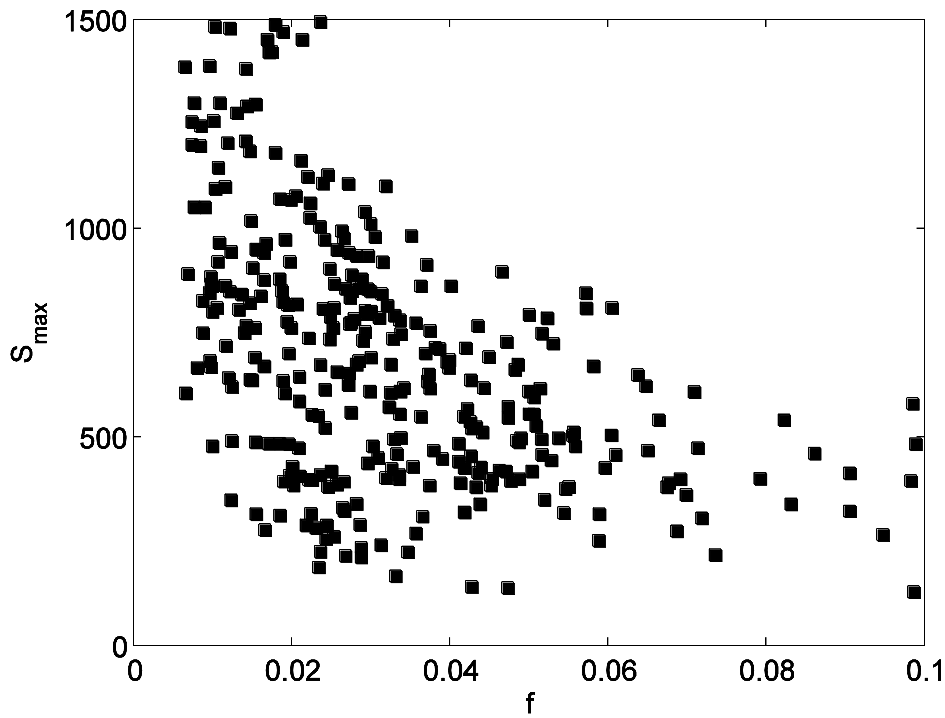

Figure 9: Relationship between calibration parameters *f* and S_{max} with data from 323 accepted catchments.

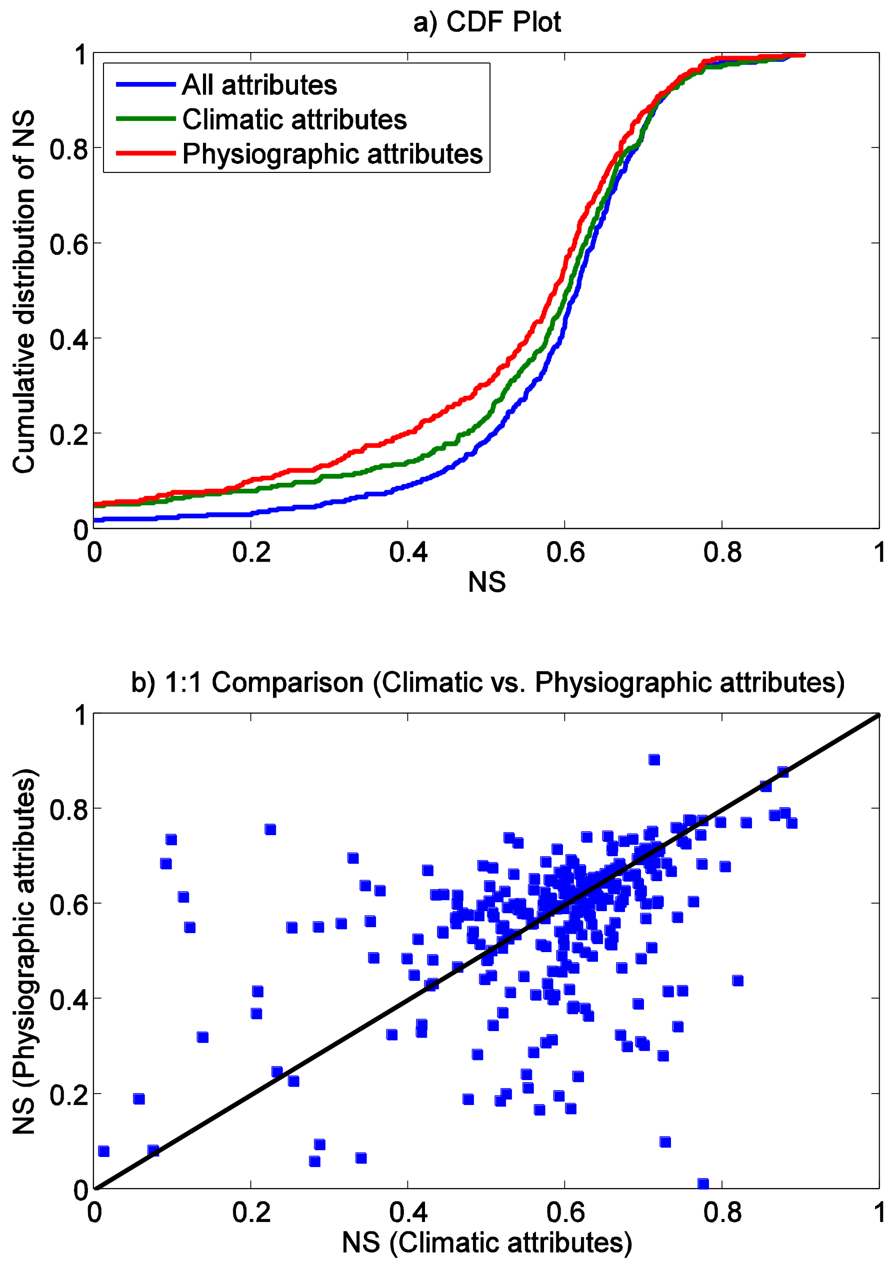

Figure 10: Comparison of climatic and physiographic attributes within the physical similarity based framework using a) CDF plot of NS values, and b) 1:1 comparison of NS values.